

Bounds on the ML Decoding Error Probability of RS-Coded Modulation over AWGN Channels

Qiutao Zhuang, Xiao Ma, *Member, IEEE*, and Aleksander Kavčić, *Member, IEEE*

Abstract

This paper is concerned with bounds on the maximum-likelihood (ML) decoding error probability of Reed-Solomon (RS) codes over additive white Gaussian noise (AWGN) channels. To resolve the difficulty caused by the dependence of the Euclidean distance spectrum on the way of signal mapping, we propose to use random mapping, resulting in an ensemble of RS-coded modulation (RS-CM) systems. For this ensemble of RS-CM systems, analytic bounds are derived, which can be evaluated from the known (symbol-level) Hamming distance spectrum. Also presented in this paper are simulation-based bounds, which are applicable to any specific RS-CM system and can be evaluated by the aid of a list decoding (in the Euclidean space) algorithm. The simulation-based bounds do not need distance spectrum and are numerically tight for short RS codes in the regime where the word error rate (WER) is not too low. Numerical comparison results are relevant in at least three aspects. First, in the short code length regime, RS-CM using BPSK modulation with random mapping has a better performance than binary random linear codes. Second, RS-CM with random mapping (time varying) can have a better performance than with specific mapping. Third, numerical results show that the recently proposed Chase-type decoding algorithm is essentially the ML decoding algorithm for short RS codes.

Q. Zhuang and X. Ma are with the Department of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou 510275, China. (Email: maxiao@mail.sysu.edu.cn)

A. Kavčić is with the Department of Electrical Engineering, University of Hawaii at Manoa, Honolulu, 96822 HI USA.

This work was supported by the 973 Program (No. 2012CB316100), by the NSF (No. 61172082) of China and by NSF grant CCF-1018984. When this work was conducted, X. Ma and A. Kavčić were visiting scholars supported by the Institute of Networking Coding at Chinese University of Hong Kong.

Index Terms

List decoding algorithm, maximum-likelihood (ML) decoding, performance bounds, Reed-Solomon (RS) codes, RS-coded modulation (RS-CM).

I. INTRODUCTION

Reed-Solomon (RS) codes are an important class of algebraic codes, which have been widely used in many practical systems, including space and satellite communications, data storage, digital audio/video transmission and file transfer [1]. The widespread use of RS codes is primarily due to their excellent error-correction capability, a consequence of their maximum distance separable (MDS) property. Hence investigating the decoding algorithms for RS codes is important in both practice and theory. The traditional hard-decision decoding (HDD) algorithms, say the Berlekamp-Massey (BM) algorithm [2], are efficient to find the unique codeword (if it exists) within a Hamming sphere of radius less than the half minimum Hamming distance. Hence, their error-correction capability is limited by the half minimum Hamming distance bound. In contrast, Guruswami-Sudan (GS) algorithm [3][4] can enlarge the decoding radius and may output a list of candidate codewords. Hence, GS algorithm can correct errors beyond the half minimum Hamming distance bound. To further improve the performance, one needs turn to the soft-decision decoding (SDD) algorithms.

The SDD algorithms with feasible complexity for RS codes include the generalized minimum distance (GMD) algorithm [5], the Chase-GMD algorithm [6], the Koetter-Vardy (KV) algorithm [7], the Chase-KV algorithm [8], the ordered statistic decoding (OSD) algorithm [9][10], and the adaptive belief propagation (ABP) algorithm [11], etc. Recently, two Chase-type decoding algorithms have been proposed for RS codes [12][13]. All these efforts have been made to improve incrementally the performance of some existed algorithms and to achieve the performance of the maximum likelihood decoding though it has been shown by Guruswami and Vardy in [14] that maximum likelihood (ML) decoding of a general RS code is NP-hard¹.

An immediate question is how to measure the sub-optimality (the gap to the ML decoding) of varieties of decoding algorithms. Though the ML decoding algorithm is prohibitively complex,

¹For short RS codes with binary phase-shift keying (BPSK) signalling, ML decoding can be performed based on the algebraic structure of their binary images [15][16].

tight bounds can be used to predict the performance without resorting to computer simulations. As mentioned in [17], most bounding techniques have connections to either the 1965 Gallager bound [18–20] or the 1961 Gallager bound [21–29] based on Gallager’s first bounding technique (GFBT). For a RS-coded modulation (RS-CM) system even with BPSK signalling, these bounds become helpless since there is no simple way to derive the bit-level Hamming weight spectrum from the known symbol-level Hamming weight distribution. The difficulty is partially caused by the dependence of the bit representation of a symbol on the choice of basis. In [16], upper bounds on ML decoding error probability were computed by computer search of the bit-level weight spectrum for short RS codes, while in [30], upper bounds were derived from the average binary weight enumerator (BWE) of the RS codes. Generally, for RS-CM using other modulations [31], the Euclidean distance spectrum of the codewords depends on the chosen signal mapping, resulting in the difficulty of performance evaluation.

At this point, we emphasize that a study of a communication system that utilizes RS codes is incomplete if only the coding aspect is considered. The performance of any communication system heavily depends on the chosen modulation and constellation mapping. It is well known that, even if the signalling constellation is fixed, the system performance still depends on which exact mapping from coded symbols to constellation points is chosen. For this reason, in this paper, our goal is to consider the analysis of the ML decoding performance of a RS code in conjunction with a high order modulation. However, incorporating a constellation mapping structure into the analysis of ML decoding of RS codes seems to only further complicate the analysis. This could be one reason why most bounds were developed for binary codes with BPSK modulation.

To resolve this difficulty, instead of considering any specific modulation mapping, we adopt a *random* mapping approach which gives rise to an *ensemble* of RS-CM systems. For such an *ensemble*, we will show that it is indeed possible to find *analytic* bounds on the performance of ML decoding error probability. *Randomization* is a powerful technique to analyze the performance, which has been widely used in the field of information and coding theory. For example, Benedetto [32] introduced *random interleaver* to derive the weight distribution of an ensemble of turbo codes, while Richardson-Urbanke [33] and Luby [34] *et al* showed how to predict the performance of LDPC codes by introducing the random irregular Tanner graphs. However, in our approach, it is not the code that is random (indeed, the code is a well-constructed algebraic

code), but it is the modulation mapping (which maps symbols to constellation points) that is random. For this ensemble of RS-CM systems, analytic bounds are derived based on the results in [29], which require only the known (symbol-level) Hamming distance spectrum of the RS code given that the signal constellation is fixed. To the best of our knowledge, a randomized approach to modulation mapping has never been used in the past as an analytic tool to study the performance of modulation mappings and this presents the first contribution of this paper.

The second contribution of this paper is that we present simulation-based bounds which are applicable to any specific RS-CM system. The simulation bounds can be evaluated by the aid of a list decoding (in the Euclidean space) algorithm. For short codes, the bounds are tight (almost overlapped). This also shows that the recently proposed tree-based Chase-type algorithm [13] is near optimal for short RS codes.

The rest of this paper is organized as follows. In Sec. II, the union bound (UB) and the sphere bound (SB) for general codes are reviewed. In Sec. III, we propose analytic bounds for the ensemble of RS-CM systems using random modulation and derive the corresponding average Euclidean distance enumerating function. In Sec. IV, we present simulation-based bounds for specific RS-CM system using the aid of a list decoding algorithm. Numerical results are presented in Sec. V, and Sec. VI concludes this paper.

II. RS-CODED MODULATION

A. System Model

Let $\mathbb{F}_q \triangleq \{\alpha_0, \alpha_1, \dots, \alpha_{q-1}\}$ be the finite field of size q . A codeword of an RS code $\mathcal{C}_q[n, k, d_{\min}]$ with length n , dimension k and minimum Hamming distance $d_{\min} = n - k + 1$ can be obtained by evaluating a polynomial of degree less than k over a set of n distinct points, denoted by $\mathcal{P} \triangleq \{\beta_0, \beta_1, \dots, \beta_{n-1}\} \subseteq \mathbb{F}_q$.

Encoding: Let $\underline{u} = (u_0, u_1, \dots, u_{k-1}) \in \mathbb{F}_q^k$ be an information sequence to be transmitted, which specifies a message polynomial $u(x) = u_0 + u_1x + \dots + u_{k-1}x^{k-1}$. The corresponding codeword is then given by

$$\underline{c} = (c_0, c_1, \dots, c_{n-1}) = (u(\beta_0), u(\beta_1), \dots, u(\beta_{n-1})). \quad (1)$$

Mapping: The codeword \underline{c} is transformed into a signal vector $\underline{s} = (s_0, s_1, \dots, s_{n-1})$, where $s_i = \phi(c_i) \in \mathbb{R}^\ell$ is an ℓ -dimensional signal which is determined by the mapping rule ϕ . The

constellation $\mathcal{X} \triangleq \{\phi(\alpha), \alpha \in \mathbb{F}_q\}$ is of size q , whose form depends on the modulation scheme. For an example, we consider a 64-ary RS code. If BPSK is implemented, we have $\ell = 6$ and $\mathcal{X} = \{-1, +1\}^6$, the six-fold Cartesian product of $\{-1, +1\}$; if 8-PSK is implemented, we have $\ell = 4$ and $\mathcal{X} = \{8\text{-PSK}\}^2$; while if 64-QAM is implemented, we have $\ell = 2$ and $\mathcal{X} = \{\pm 1, \pm 3, \pm 5, \pm 7\}^2$. All signal vectors are collectively denoted by $\mathcal{S} \triangleq \{\underline{s} \mid s_t = \phi(c_t), 0 \leq t \leq n-1, \underline{c} \in \mathcal{C}_q\}$.

Channel: Assume that the signal vector \underline{s} is transmitted through an AWGN channel. The received vector is denoted by $\underline{y} = \underline{s} + \underline{z}$, where \underline{z} is a sample from a white Gaussian noise process with zero mean and double-sided power spectral density σ^2 .

ML Decoding: Assume that each codeword is transmitted with equal probability. The optimal decoding that minimizes the word-error probability is the ML decoding, which, for AWGN channels, is equivalent to finding the nearest signal vector $\hat{\underline{s}} \in \mathcal{S}$ to \underline{y} .

Hereafter, we may not distinguish \underline{c} from \underline{s} when representing a codeword of RS-CM.

B. Distance Enumerating Functions of RS-CM

The weight enumerating function of a RS code $\mathcal{C}_q[n, k, d_{\min}]$ is defined as

$$W(X) \triangleq \sum_{i=d_{\min}}^n W_i X^i, \quad (2)$$

where X is a dummy variable and W_i denotes the number of codewords having Hamming weight i , which can be determined by [35]

$$W_i = \binom{n}{i} (q-1) \sum_{j=0}^{i-d_{\min}} (-1)^j \binom{i-1}{j} q^{i-j-d_{\min}}, i \geq d_{\min}. \quad (3)$$

Given a codeword \underline{s} of RS-CM, we denote $A_{\delta|\underline{s}}$ the number of codewords having the Euclidean distance δ with \underline{s} . We define

$$A_\delta = \frac{1}{q^k} \sum_{\underline{s}} A_{\delta|\underline{s}},$$

which is the average number of *ordered* pairs of codewords with Euclidean distance δ .

Definition 1: The *Euclidean distance enumerating function* of RS-CM is defined as [29, (2)]

$$A(X) \triangleq \sum_{\delta} A_\delta X^{\delta^2}, \quad (4)$$

where X is a dummy variable and the summation is over all possible distance δ . We call $\{A_\delta\}$ the Euclidean distance spectrum.

C. Upper Bounds for RS-CM

The conventional union bound (UB) on the ML decoding error probability $\Pr\{E\}$ for RS-CM is

$$\Pr\{E\} \leq \sum_{\delta} A_{\delta} Q\left(\frac{\delta}{2\sigma}\right), \quad (5)$$

where $Q\left(\frac{\delta}{2\sigma}\right)$ is the pair-wise error probability with

$$Q(x) \triangleq \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz. \quad (6)$$

The UB can be tightened by the use of the following sphere bound (SB) as shown in [29, (26)],

$$\Pr\{E\} \leq \int_0^{+\infty} \min\{f_u(r), 1\} g(r) dr, \quad (7)$$

where

$$f_u(r) = \sum_{\delta} A_{\delta} p_2(r, \delta), \quad (8)$$

$$p_2(r, \delta) = \begin{cases} \frac{\Gamma(\frac{\ell n}{2})}{\sqrt{\pi} \Gamma(\frac{\ell n-1}{2})} \int_0^{\arccos(\frac{\delta}{2r})} \sin^{\ell n-2} \phi d\phi, & r > \frac{\delta}{2} \\ 0, & r \leq \frac{\delta}{2} \end{cases}, \quad (9)$$

and

$$g(r) = \frac{2r^{\ell n-1} e^{-\frac{r^2}{2\sigma^2}}}{2^{\frac{\ell n}{2}} \sigma^{\ell n} \Gamma(\frac{\ell n}{2})}, \quad r \geq 0, \quad (10)$$

which is determined by the Euclidean distance spectrum $\{A_{\delta}\}$.

III. ANALYTIC BOUNDS FOR AN ENSEMBLE OF RS-CODED MODULATION SYSTEMS

As seen from Sec. II, computing the derived upper bounds on the ML decoding error probability for RS-CM requires the Euclidean distance spectrum $\{A_{\delta}\}$, which depends on the way of signal mapping ϕ and is usually difficult to compute. To resolve this difficulty, in this section, we propose to use random mapping, resulting in an ensemble of RS-CM systems. For this ensemble of RS-CM systems, analytic bounds are derived, which can be evaluated from the weight enumerating function $W(X)$.

A. Average Euclidean Distance Enumerating Function of RS-CM with Random Modulation

Let $\mathcal{C}_q[n, k, d_{\min}]$ be an RS code defined over \mathbb{F}_q and $\mathcal{X} \subset \mathbb{R}^\ell$ be a signal constellation of size q . Let $\Phi = \{\phi^{(1)}, \phi^{(2)}, \dots, \phi^{(q!)}\}$ be the set of all one-to-one mapping rules from \mathbb{F}_q to \mathcal{X} . Assume that $\underline{\phi} = (\phi_0, \phi_1, \dots, \phi_{n-1})$ is a random sequence, whose components are sampled independently and uniformly from Φ . Define $\mathcal{S}(\mathcal{C}_q, \underline{\phi}) \triangleq \{\underline{s} \mid s_t = \phi_t(c_t), 0 \leq t \leq n-1, \underline{c} \in \mathcal{C}_q\}$. It can be seen that $\mathcal{S}(\mathcal{C}_q, \underline{\phi}) \subset \mathbb{R}^{\ell n}$ is a random codebook of size q^k . Thus the RS code \mathcal{C}_q can be mapped to $(q!)^n$ different codebooks \mathcal{S} , each of which with probability $\Pr\{\mathcal{S}\} = 1/(q!)^n$.

Given a codebook $\mathcal{S}(\mathcal{C}_q, \underline{\phi})$, we denote $B_\delta(\mathcal{S})$ the average number of *ordered* pairs of codewords with Euclidean distance δ in $\mathcal{S}(\mathcal{C}_q, \underline{\phi})$. We define

$$B_\delta \triangleq \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} B_\delta(\mathcal{S}), \quad (11)$$

which denotes the ensemble average of the number of ordered pairs of codewords with Euclidean distance δ .

Definition 2: The *average Euclidean distance enumerating function* of RS-CM with random modulation is defined as

$$B(X) \triangleq \sum_{\delta} B_\delta X^{\delta^2}. \quad (12)$$

We call $\{B_\delta\}$ the average Euclidean distance spectrum.

B. Analytic Bounds for the Ensemble of RS-CM with Random Modulation

The ML decoding error probability $\Pr\{E\}$ for the ensemble of RS-CM can be written as

$$\Pr\{E\} = \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} \Pr\{E|\mathcal{S}\}, \quad (13)$$

where $\Pr\{E|\mathcal{S}\}$ is the conditional ML decoding error probability given a code $\mathcal{S}(\mathcal{C}_q, \underline{\phi})$.

From (5), the UB of $\Pr\{E|\mathcal{S}\}$ is

$$\Pr\{E|\mathcal{S}\} \leq \sum_{\delta} B_\delta(\mathcal{S}) Q\left(\frac{\delta}{2\sigma}\right). \quad (14)$$

Therefore, from (11) and (13), the UB on the ML decoding error probability of the ensemble

of RS-CM can be written as

$$\begin{aligned}
\Pr\{E\} &= \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} \Pr\{E|\mathcal{S}\} \\
&\leq \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} \sum_{\delta} B_{\delta}(\mathcal{S}) Q\left(\frac{\delta}{2\sigma}\right) \\
&= \sum_{\delta} \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} B_{\delta}(\mathcal{S}) Q\left(\frac{\delta}{2\sigma}\right) \\
&= \sum_{\delta} B_{\delta} Q\left(\frac{\delta}{2\sigma}\right), \tag{15}
\end{aligned}$$

which is determined by the average Euclidean distance spectrum $\{B_{\delta}\}$.

From (7), the SB of $\Pr\{E|\mathcal{S}\}$ is

$$\Pr\{E|\mathcal{S}\} \leq \int_0^{+\infty} \min\{f_u(r|\mathcal{S}), 1\} g(r) \, dr, \tag{16}$$

where $g(r)$ is given by (10) and $f_u(r|\mathcal{S}) = \sum_{\delta} B_{\delta}(\mathcal{S}) p_2(r, \delta)$ with $p_2(r, \delta)$ given by (9).

From (11), we define

$$\begin{aligned}
f_u(r) &\triangleq \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} f_u(r|\mathcal{S}) \\
&= \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} \sum_{\delta} B_{\delta}(\mathcal{S}) p_2(r, \delta) \\
&= \sum_{\delta} \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} B_{\delta}(\mathcal{S}) p_2(r, \delta) \\
&= \sum_{\delta} B_{\delta} p_2(r, \delta). \tag{17}
\end{aligned}$$

Therefore, from (13) and (17), the SB on the ML decoding error probability of the ensemble of RS-CM can be written as

$$\begin{aligned}
\Pr\{E\} &= \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} \Pr\{E|\mathcal{S}\} \\
&\leq \sum_{\mathcal{S}} \Pr\{\mathcal{S}\} \int_0^{+\infty} \min\{f_u(r|\mathcal{S}), 1\} g(r) \, dr \\
&\leq \int_0^{+\infty} \min\left\{\sum_{\mathcal{S}} \Pr\{\mathcal{S}\} f_u(r|\mathcal{S}), 1\right\} g(r) \, dr \\
&\leq \int_0^{+\infty} \min\{f_u(r), 1\} g(r) \, dr, \tag{18}
\end{aligned}$$

which is determined by the average Euclidean distance spectrum $\{B_{\delta}\}$.

C. Computation of the Average Euclidean Distance Enumerating Function

As seen from the preceding subsection, computing the derived bounds for the ensemble of RS-CM requires the average Euclidean distance enumerating function $B(X)$ defined in (12). In this subsection, we will show that $B(X)$ is computable from the Hamming weight enumerating function $W(X)$ and the Euclidean distance enumerating function of the signal constellation.

Recall that \mathcal{X} is the signal constellation of size q . Define

$$D(X) \triangleq \sum_{\delta} D_{\delta} X^{\delta^2} = \frac{1}{q(q-1)} \sum_{x \in \mathcal{X}, y \in \mathcal{X}, x \neq y} X^{\|x-y\|^2}, \quad (19)$$

where $\|\cdot\|$ denotes the Euclidean distance and D_{δ} denotes the average number of *ordered* signal pairs with Euclidean distance δ in the constellation \mathcal{X} . Let \underline{c} and $\hat{\underline{c}}$ be two codewords with Hamming distance $d > 0$. Denote by $\underline{\phi} = (\phi_0, \phi_1, \dots, \phi_{n-1})$ the sequence of random mappings, whose components are independent and uniformly distributed over Φ , the set of all one-to-one mappings. Then we have,

$$\sum_{\underline{\phi}} (q!)^{-n} X^{\sum_t \|\phi_t(c_t) - \phi_t(\hat{c}_t)\|^2} = (D(X))^d.$$

Therefore,

$$\begin{aligned} B(X) &= \sum_{\underline{\phi}} (q!)^{-n} \sum_{\underline{c} \in \mathcal{C}_q} q^{-k} \sum_{\hat{\underline{c}} \in \mathcal{C}_q, \hat{\underline{c}} \neq \underline{c}} X^{\sum_t \|\phi_t(c_t) - \phi_t(\hat{c}_t)\|^2} \\ &= \sum_{\underline{c} \in \mathcal{C}_q} q^{-k} \sum_{\hat{\underline{c}} \in \mathcal{C}_q, \hat{\underline{c}} \neq \underline{c}} \sum_{\underline{\phi}} (q!)^{-n} X^{\sum_t \|\phi_t(c_t) - \phi_t(\hat{c}_t)\|^2} \\ &= \sum_{d_{\min} \leq d \leq n} W_d \cdot (D(X))^d. \end{aligned}$$

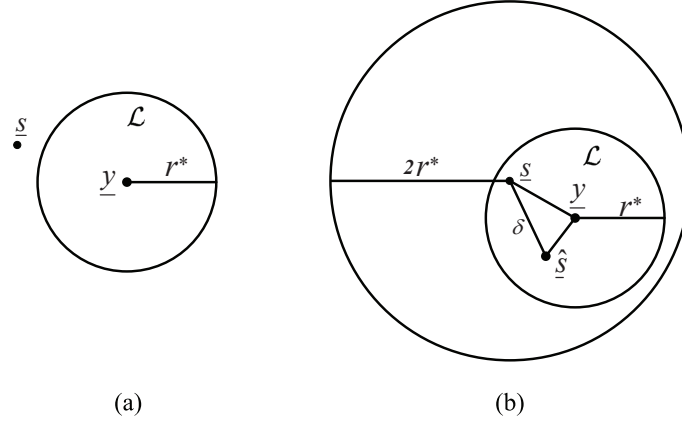
Remark. When considering RS-CM using BPSK modulation, the average Euclidean distance enumerating function $B(X)$ is reduced to the average binary weight enumerator (BWE) presented in [30].

IV. SIMULATION-BASED BOUNDS FOR SPECIFIC RS-CODED MODULATION SYSTEM

In this section, we present a simulation-based bound for specific RS-CM system by the aid of a list decoding algorithm.

Algorithm 1: A suboptimal list decoding algorithm for the purpose of performance analysis

- S1. List all codewords within the Euclidean sphere with center at the received vector \underline{y} and with radius $r^* \geq 0$ where r^* is a parameter to be determined. We denote the list as \mathcal{L} .



(a) The error event that the transmitted codeword is not in the list.
(b) The error event that the transmitted codeword is in the list but not the closest one.

Fig. 1. Graphical illustrations of the decoding error events.

S2. Find the codeword $\underline{s}^* \in \mathcal{L}$ that is closest to \underline{y} .

Remark. The above list decoding algorithm was also referred to as sphere decoding algorithm in [36]. The objective of [36] is to derive analytic bounds on the sphere decoding itself, while our objective here is to derive simulation-based bounds on the ML decoding by assuming that the sphere decoding can be implemented efficiently.

The decoding error occurs in two cases under the assumption that the codeword \underline{s} is transmitted.

Case 1. The transmitted codeword \underline{s} is not in the list \mathcal{L} (see Fig. 1 (a)), that is, $\|\underline{z}\| = \|\underline{y} - \underline{s}\| \geq r^*$. This event is denoted by $\{E_1|\underline{s}\}$, whose probability is given by

$$\begin{aligned} \Pr\{E_1|\underline{s}\} &= \Pr\{\|\underline{z}\| \geq r^*\} \\ &= \int_{r^*}^{+\infty} g(r) \, dr, \end{aligned} \quad (20)$$

where $g(r)$ is defined in (10).

Case 2. The transmitted codeword \underline{s} is in the list \mathcal{L} , but is not the closest one (see Fig. 1 (b)). The event is denoted by $\{E_2|\underline{s}\}$, resulting in the error probability $\Pr\{E_2|\underline{s}\}$.

Obviously, $\Pr\{E_2|\underline{s}\} \leq \Pr\{E|\underline{s}\} \leq \Pr\{E_1|\underline{s}\} + \Pr\{E_2|\underline{s}\}$. Averaging over the transmitted codewords, we have

$$\Pr\{E_2\} \leq \Pr\{E\} \leq \Pr\{E_1\} + \Pr\{E_2\}, \quad (21)$$

where $\Pr\{E_1\} = \sum_{\underline{s}} \Pr\{\underline{s}\} \Pr\{E_1|\underline{s}\} = \int_{r^*}^{+\infty} g(r) \, dr$ and $\Pr\{E_2\} = \sum_{\underline{s}} \Pr\{\underline{s}\} \Pr\{E_2|\underline{s}\}$. If we can calculate $\Pr\{E_2\}$, then we have bounds on the ML decoding error probability. Actually, for small r^* or short RS codes, this probability $\Pr\{E_2\}$ can be estimated by *Monte Carlo* simulation using the recently proposed tree-based Chase-type algorithm [13].

Algorithm 2: Estimate the error probability $\Pr\{E_2\}$.

```

1: Initialize  $i = 0$  and  $N_{err} = 0$ . Given a parameter  $r^* > 0$  and a sufficiently large integer
    $N_{total} > 0$ .
2: while ( $i < N_{total}$ ) do
3:   Generate uniformly at random a codeword  $\underline{s}$  and a white Gaussian noise sample  $\underline{z}$ .
4:    $\underline{y} \leftarrow \underline{s} + \underline{z}$ 
5:   if  $\|\underline{y} - \underline{s}\| \leq r^*$  then
6:      $i \leftarrow i + 1$ 
7:     Decode  $\underline{y}$  with the algorithm [13], resulting in the decoded codeword  $\underline{s}^*$ .
8:     if  $\underline{s}^*$  is different from  $\underline{s}$  then
9:        $N_{err} \leftarrow N_{err} + 1$ 
10:    end if
11:  end if
12: end while
13:  $\Pr\{E_2\} = N_{err}/N_{total}$ 
14: return  $\Pr\{E_2\}$ 

```

Remark. Notice that the above algorithm can be faster than the real decoding algorithm since we do not have to find the optimal candidate codeword within the sphere in the case that some *earlier* intermediate candidate is found to be better than the transmitted one².

V. NUMERICAL RESULTS

Example 1: In this example, we consider $\mathcal{C}_{16}[15, 11, 5]$ using BPSK modulation. For comparison, we also consider a random binary linear code $\mathcal{C}_2[60, 44]$ with BPSK modulation. Both these two codes have the same code length and code rate. The numerical results are shown in Fig. 2. We can see that the SB is tighter than the UB in the low signal-to-noise ratio (SNR) regime, as

²In this case, the ML decoding must also make an error, as pointed out in [37] and used in [38].

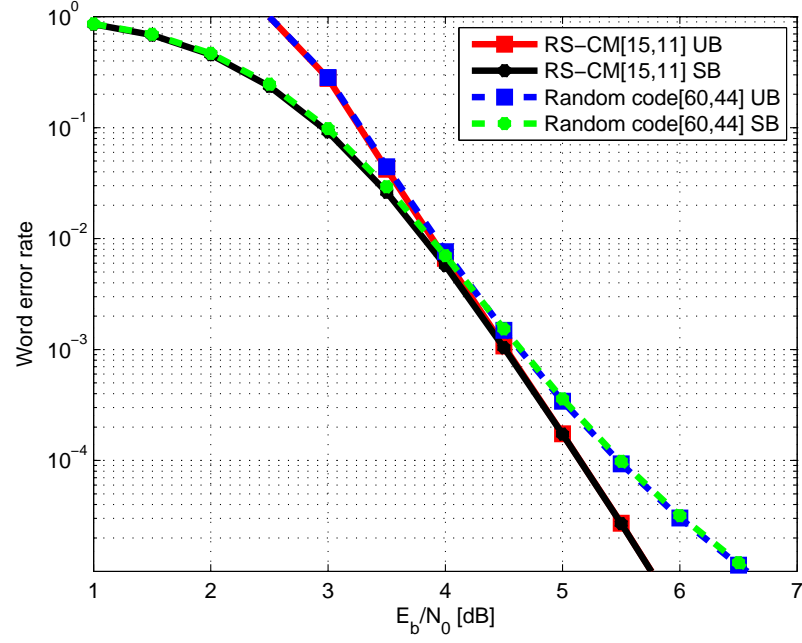


Fig. 2. Upper bounds on the word error rate (WER) of the ensemble of $\mathcal{C}_{16}[15, 11, 5]$ RS-CM using BPSK modulation.

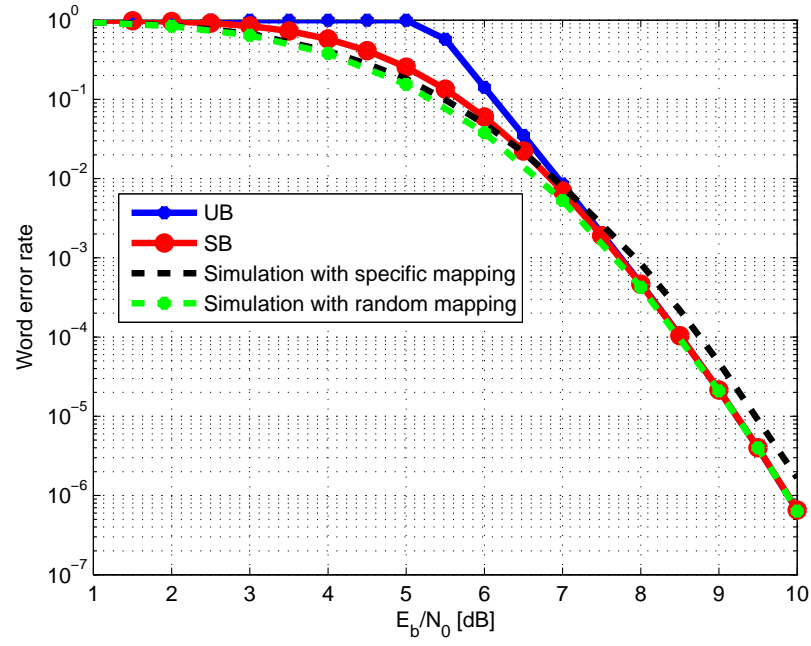


Fig. 3. Upper bounds on the WER of the ensemble of $\mathcal{C}_{16}[15, 11, 5]$ RS-CM using 16-QAM signal constellation.

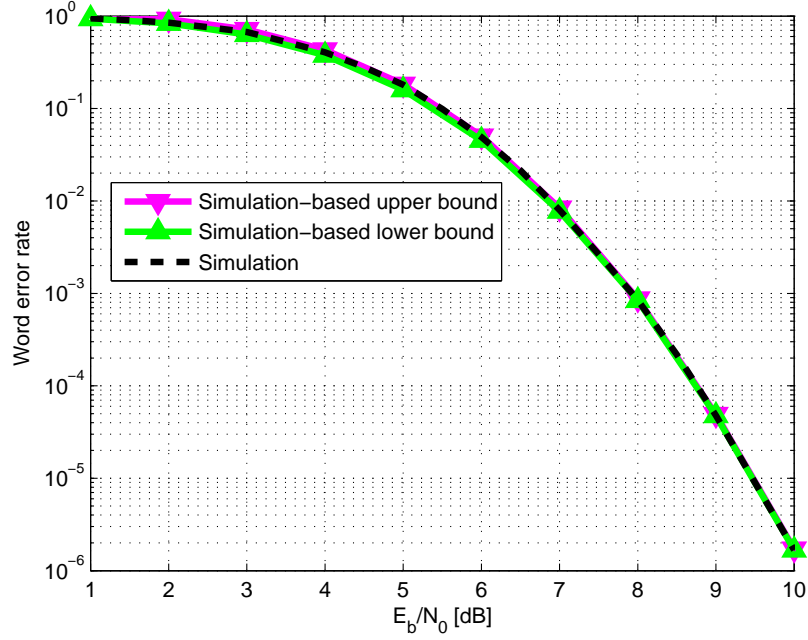


Fig. 4. Simulation-based bounds on the WER of $C_{16}[15, 11, 5]$ RS-CM using 16-QAM signal constellation.

expected. The results also show that this nonbinary structured code is better (on average) than binary random linear codes. Since binary linear codes can achieve the capacity as the code length goes to infinity, this numerical result indicates that, in short-length regime, the code structure plays a more important role.

Example 2: In this example, we consider the same code as in Example 1 but using 16-QAM modulation. The numerical results are shown in Fig. 3. We can see that the SB is tighter than the UB in the low SNR regime. We can also see that, in the high SNR regime, we do not have to use the complicated SB to predict the performance. Also presented in Fig. 3 are the simulation results with specific/random modulation using the recently proposed tree-based Chase-type algorithm [13]. It can be seen that the simulated curve with random mapping matches well with the analytic bounds for the ensemble and the simulated curve with specific mapping is slightly worse than the ensemble upper bounds in the high SNR regime. This comparison also shows that the performance of RS-CM is closely related to which modulation is chosen.

Example 3: In this example, we consider the same RS-CM system as in Example 2 but using specific modulation. The numerical results are shown in Fig. 4. Also presented in Fig. 4 are the

simulation results using the same decoding algorithm [13] as in Example 2. It can be seen that all these curves are almost overlapped, showing that the decoding algorithm is near optimal.

VI. CONCLUSIONS

In this paper, we have presented analytic bounds for the ensemble of RS-CM systems using random modulation. We have also presented simulation-based bounds for any specific RS-CM systems. Numerical results show that, at least in the short code-length regime, the performance of RS-CM with random modulation is better than that of random linear codes. Numerical results also show that the recently proposed Chase-type decoding algorithm is near optimal.

ACKNOWLEDGMENT

The authors would like to thank Siyun Tang for helpful discussions in connection to the simulation-based bounds.

REFERENCES

- [1] D. J. Costello, Jr., J. Hagenauer, H. Imai, and S. B. Wicker, "Applications of error-control coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2531–2560, Oct. 1998.
- [2] E. R. Berlekamp, "Nonbinary BCH decoding," *IEEE Trans. Inf. Theory*, vol. IT-14, p. 242, Mar. 1968.
- [3] M. Sudan, "Decoding of Reed Solomon codes beyond the error-correction bound," *Journal of Complexity*, vol. 13, no. 1, pp. 180–193, Mar. 1997.
- [4] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon and algebraic-geometric codes," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1757–1767, Sept. 1999.
- [5] G. D. Forney, Jr., "Generalized minimum distance decoding," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 2, pp. 125–131, Apr. 1966.
- [6] H. Tang, Y. Liu, M. Fossorier, and S. Lin, "On combining Chase-2 and GMD decoding algorithms for nonbinary block codes," *IEEE Commun. Lett.*, vol. CL-5, pp. 209–211, May 2001.
- [7] R. Koetter and A. Vardy, "Algebraic soft-decision decoding of Reed Solomon codes," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2809–2825, Nov. 2003.
- [8] X. Zhang, Y. Zheng, and Y. Wu, "A Chase-type Koetter-Vardy algorithm for soft-decision Reed-Solomon decoding," in *Proc. 2012 Int. Conf. Comput., Netw. Commun.*, pp. 466–470.
- [9] M. P. Fossorier and S. Lin, "Soft-decision decoding of linear block codes based on ordered statistics," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1379–1396, Sept. 1995.
- [10] W. Jin and M. P. Fossorier, "Towards maximum likelihood soft decision decoding of the (255,239) Reed Solomon code," *IEEE Trans. Magn.*, vol. 44, no. 3, pp. 423–428, Mar. 2008.

- [11] J. Jiang and K. R. Narayanan, "Iterative soft-input soft-output decoding of Reed-Solomon codes by adapting the parity-check matrix," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3746–3756, Aug. 2006.
- [12] J. Bellorado and A. Kavčić, "Low-complexity soft-decoding algorithms for Reed-Solomon codes-part I: An algebraic soft-in hard-out Chase decoder," *IEEE Trans. Inf. Theory*, vol. 56, no. 3, pp. 945–959, Mar. 2010.
- [13] S. Tang and X. Ma, "A new Chase-type soft-decision decoding algorithm for Reed-Solomon codes," submitted to *IEEE Trans. Inf. Theory*. [Online]. Available: <http://arxiv.org/abs/1309.1555>
- [14] V. Guruswami and A. Vardy, "Maximum-likelihood decoding of Reed-Solomon codes is NP-hard," *IEEE Trans. Inf. Theory*, vol. 51, pp. 2249–2256, July 2005.
- [15] A. Vardy and Y. Be'ery, "Bit-level soft-decision decoding of Reed-Solomon codes," *IEEE Trans. Commun.*, vol. 39, no. 3, pp. 440–444, Mar. 1991.
- [16] V. Ponnampalam and B. Vucetic, "Soft decision decoding for Reed-Solomon codes," *IEEE Trans. Commun.*, vol. 50, no. 11, pp. 1758–1768, Nov. 2002.
- [17] I. Sason and S. Shamai, "Performance analysis of linear codes under maximum-likelihood decoding: A tutorial," in *Foundations and Trends in Commun. and Inf. Theory*. Delft, The Netherlands: NOW, July 2006, vol. 3, no. 1-2, pp. 1–225.
- [18] T. M. Duman and M. Salehi, "New performance bounds for turbo codes," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 717–723, June 1998.
- [19] T. M. Duman, "Turbo codes and turbo coded modulation systems: Analysis and performance bounds," Ph.D. dissertation, Elect. Comput. Eng. Dept., Northeastern Univ., Boston, MA, May 1998.
- [20] M. Twitto, I. Sason, and S. Shamai, "Tightened upper bounds on the ML decoding error probability of binary linear block codes," *IEEE Trans. Inf. Theory*, vol. 53, pp. 1495–1510, Apr. 2007.
- [21] H. Herzberg and G. Poltyrev, "Techniques of bounding the probability of decoding error for block coded modulation structures," *IEEE Trans. Inf. Theory*, vol. 40, pp. 903–911, May 1994.
- [22] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1284–1292, July 1994.
- [23] D. Divsalar, "A simple tight bound on error probability of block codes with application to turbo codes," in *Proc. 1999 IEEE Commun. Theory Workshop*, Aptos, CA, May 1999.
- [24] D. Divsalar and E. Biglieri, "Upper bounds to error probabilities of coded systems beyond the cutoff rate," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2011–2018, Dec. 2003.
- [25] S. Yousefi and A. K. Khandani, "A new upper bound on the ML decoding error probability of linear binary block codes in AWGN interference," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3026–3036, Nov. 2004.
- [26] —, "Generalized tangential sphere bound on the ML decoding error probability of linear binary block codes in AWGN interference," *IEEE Trans. Inf. Theory*, vol. 50, pp. 2810–2815, Nov. 2004.
- [27] A. Mehrabian and S. Yousefi, "Improved tangential sphere bound on the ML decoding error probability of linear binary block codes in AWGN and block fading channels," *IEE Proc. Commun.*, vol. 153, pp. 885–893,

Dec. 2006.

- [28] X. Ma, J. Liu, and B. Bai, “New techniques for upper-bounding the ML decoding performance of binary linear codes,” *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 842–851, Mar. 2013.
- [29] Q. Zhuang, J. Liu, and X. Ma, “Upper bounds on the ML decoding error probability for general codes over AWGN channels,” submitted to *IEEE Trans. Inf. Theory*. [Online]. Available: <http://arxiv.org/abs/1308.3303>
- [30] M. El-Khamy and R. J. McEliece, “Bounds on the average binary minimum distance and the maximum likelihood performance of Reed Solomon codes,” *42nd Allerton Conf. on Communication, Control and Computing*, 2004.
- [31] K.-P. Yar, D.-S. Yoo, and W. Stark, “Performance of RS coded M-ary modulation with and without symbol overlapping,” *IEEE Trans. Commun.*, vol. 56, no. 3, pp. 445–453, Mar. 2008.
- [32] S. Benedetto and G. Montorsi, “Unveiling turbo codes: Some results on parallel concatenated coding schemes,” *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 409–428, Mar. 1996.
- [33] T. Richardson and R. Urbanke, “The capacity of low-density parity check codes under message-passing decoding,” *IEEE Trans. Inf. Theory*, vol. 47, pp. 599–618, Feb. 2001.
- [34] M. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. Spielman, “Analysis of low density codes and improved designs using irregular graphs,” in *Proc. 30th Annu. ACM Symp. Theory of Computing*, pp. 249–258, 1998.
- [35] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005.
- [36] M. El-Khamy, H. Vikalo, B. Hassibi, and R. J. McEliece, “Performance of sphere decoding of block codes,” *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2940–2950, Oct. 2009.
- [37] B. G. Dorsch, “A decoding algorithm for binary block codes and J-ary output channels,” *IEEE Trans. Inf. Theory*, vol. IT-20, pp. 391–394, May 1974.
- [38] A. Valembois and M. Fossorier, “Box and match techniques applied to soft-decision decoding,” *IEEE Trans. Inf. Theory*, vol. 50, pp. 796–810, May 2004.